

⑬ BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑫ Patentschrift
⑪ DE 37 42 929 C 1

⑤① Int. Cl. 4:
G 10 L 7/08
G 10 L 3/00
H 04 B 15/02

⑳ Aktenzeichen: P 37 42 929.9-53
㉑ Anmeldetag: 18. 12. 87
㉒ Offenlegungstag: —
㉓ Veröffentlichungstag
der Patenterteilung: 29. 9. 88

DE 37 42 929 C 1

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden

⑦③ Patentinhaber:
Daimler-Benz AG, 7000 Stuttgart, DE

⑦② Erfinder:
Gollmar, Klaus, Dipl.-Ing., 7240 Horb, DE; Jünemann,
Gerhard, Dr.rer.nat., 7770 Überlingen, DE

⑤② Für die Beurteilung der Patentfähigkeit
in Betracht gezogene Druckschriften:

DE	36 10 797 A1
DE	34 01 883 A1
GB	20 03 002 A
US	33 83 466

⑤④ Verfahren zur Verbesserung der Zuverlässigkeit von Sprachsteuerungen von Funktionselementen und Vorrichtung zu dessen Durchführung

Bei einem Verfahren und einer Vorrichtung zur sicheren Erkennung mindestens des Beginns eines Sprachbefehls für eine Sprachsteuerung von Funktionselementen werden Signale eines mit Körperschall einer Bedienperson beaufschlagten ersten Mikrofons zur Triggerung eines auf den Mund der Bedienperson ausgerichteten zweiten Mikrofons verwendet, um die Zuverlässigkeit der Sprachsteuerung bei Nebengeräuschbelastung zu verbessern.

DE 37 42 929 C 1

Patentansprüche

1. Verfahren zur Verbesserung der Zuverlässigkeit von Sprachsteuerungen von Funktionselementen an Arbeitsplätzen oder in Kraftfahrzeugen, durch Vorgehens-
5 Vorkehrungen zur sicheren Erkennung des Beginns von Sprachbefehlen in einer Spracherkennungsschaltung,
wobei ein erstes Mikrofon vornehmlich mit beim Sprechen erzeugtem Schall beaufschlagt wird
10 und ein zweites Mikrofon wenigstens mit Nebengeräuschen beaufschlagt wird
und wobei in beiden Mikrofonen erzeugte Signale elektrisch in eine bestimmte Beziehung zueinander gebracht werden, um eine fälschliche Interpretation eines reinen Nebengeräuschs als Sprachbefehl zu verhindern,
15 dadurch gekennzeichnet,
daß das erste Mikrofon mit beim Sprechen erzeugtem Körperschall einer Bedienperson beaufschlagt wird und
20 daß das zweite Mikrofon, welches mit räumlichem Abstand auf den Mund der Bedienperson ausgerichtet ist und dadurch ferner mit beim Sprechen erzeugtem Schall beaufschlagt wird, nur dann elektrisch zur Sprachsignalübertragung auf die Spracherkennungsschaltung geschaltet wird, wenn das erste Mikrofon aus dem Körperschall seinerseits Signale erzeugt.
2. Vorrichtung zur Durchführung des Verfahrens nach Anspruch 1
30 mit einem ersten, vornehmlich mit beim Sprechen erzeugten Schallwellen beaufschlagbaren Mikrofon,
mit einem zweiten, wenigstens mit Nebengeräuschen beaufschlagbaren Mikrofon und
35 mit einer elektrischen Schaltung, welche eine bestimmte Beziehung zwischen von den beiden Mikrofonen erzeugten elektrischen Signalen herstellt und einer Spracherkennungsschaltung Ausgangssignale zuführt,
40 dadurch gekennzeichnet,
daß das erste Mikrofon (2) als an einem Körperteil einer Bedienperson anliegendes Körperschallmikrofon ausgebildet ist,
45 daß das zweite Mikrofon (5) mit räumlichem Abstand auf den Mund der Bedienperson ausgerichtet ist und ferner mit beim Sprechen erzeugtem Schall beaufschlagt wird, und
daß die elektrische Schaltung als durch vom ersten Mikrofon (2) erzeugte Signale ansteuerbare Triggerschaltung (10) zur Herstellung einer Schaltbeziehung zwischen Signalen des ersten (2) und des zweiten Mikrofonen (5) ausgebildet ist, welche
50 Schaltbeziehung eine Zuführung von Signalen mindestens des zweiten Mikrofonen (5) als Ausgangssignale zur Spracherkennungsschaltung (14) nur dann ermöglicht, wenn gleichzeitig Signale des ersten Mikrofonen (2) an der Triggerschaltung (10) anliegen.
3. Vorrichtung nach Anspruch 2, dadurch gekennzeichnet, daß das erste Mikrofon (2) ein an den Hals der Bedienperson anlegbares Kehlkopfmikrofon ist.
4. Vorrichtung nach Anspruch 3, dadurch gekennzeichnet, daß beide Mikrofone (2, 5) an Geschirren (1, 3) befestigt sind, die von der Bedienperson angelegt werden können.

5. Vorrichtung nach Anspruch 2, dadurch gekennzeichnet, daß ferner ein von der Bedienperson zu tragendes Gehäuse (9) zur Aufnahme

- der mit beiden Mikrofonen (2, 5) elektrisch verbundenen Triggerschaltung (10),
- eines Senders (11) zur drahtlosen Übertragung von Ausgangssignalen der Triggerschaltung (10) an einen Empfänger (13) der Spracherkennungsschaltung (14) und
- eines elektrischen Energiespeichers (12) zur Speisung mindestens der Triggerschaltung (10) und des Senders (11)

vorgesehen ist.

6. Vorrichtung nach Anspruch 2, dadurch gekennzeichnet, daß die Signale des ersten Mikrofonen (2) ferner zur Optimierung der Erkennbarkeit der Signale des zweiten Mikrofonen (5) mit diesen elektrisch verknüpft werden.

7. Vorrichtung nach Anspruch 2, dadurch gekennzeichnet, daß ein Zeitglied (7) mit auf die maximale Wortempfangsdauer der Spracherkennungsschaltung (14) einstellbarer Zeitkonstante vorgesehen ist, welches die Zuführung von Ausgangssignalen der Triggerschaltung (10) zur Spracherkennungsschaltung nach Ablauf der eingestellten Zeitkonstante unterbricht.

8. Vorrichtung nach Anspruch 7, dadurch gekennzeichnet, daß ein Signalgeber (10.3; 4) vorgesehen ist, welcher beim Ablauf der Zeitkonstante zur Erzeugung eines Signals an die Bedienperson aktiviert wird.

9. Vorrichtung nach Anspruch 2, dadurch gekennzeichnet,

daß die Triggerschaltung (10) einen Schwellwertschalter (10.1) aufweist, welcher die elektrischen Signale des ersten Mikrofonen (2) mit einem einstellbaren Schwellwert vergleicht und

daß ein von dem Schwellwertschalter (10.1) bei über dem Schwellwert liegenden Signalen des ersten Mikrofonen (2) schließbarer Schalter (10.2) in den Weg der elektrischen Signale des zweiten Mikrofonen (5) eingeschleift ist.

10. Vorrichtung nach Anspruch 5, dadurch gekennzeichnet,

daß das Gehäuse (9) einen ersten Gehäuseteil (9.1) zur Aufnahme der Triggerschaltung (10) und einen zweiten Gehäuseteil (9.2) zur Aufnahme des Senders (11) und des Energiespeichers (12) aufweist und

daß die Triggerschaltung (10) mit dem Sender (11) und dem Energiespeicher (12) elektrisch mittels einer zwischen den beiden Gehäuseteilen (9.1, 9.2) angeordneten und diese kuppelnden mehrpoligen Steckverbindung (15; Fig. 2) verbunden ist.

11. Vorrichtung nach Anspruch 9, dadurch gekennzeichnet, daß das Zeitglied (7) in den elektrischen Schalter (10.2) integriert ist.

12. Vorrichtung nach Anspruch 2, dadurch gekennzeichnet, daß in eine Verbindung (6) zwischen dem ersten Mikrofon (2) und der Triggerschaltung (10) ein manuell zu betätigender Unterbrecherschalter (6.1) eingeschleift ist.

Beschreibung

Die Erfindung bezieht sich auf ein Verfahren nach

dem Oberbegriff des Verfahrensanspruchs 1 und auf eine Vorrichtung zur Durchführung des erfindungsgemäßen Verfahrens, welche die Merkmale des Oberbegriffs des Vorrichtungsanspruchs 2 aufweist.

Ein gattungsgemäßes Verfahren läßt sich aus der Arbeitsweise einer aus der DE-OS 36 10 797 bekannten Vorrichtung herleiten. Diese enthält bereits ein Sprachmikrofon und ein Nebengeräuschmikrofon, deren Signale durch elektrische Subtraktion zueinander in Beziehung gesetzt werden. Bei der Subtraktion werden die beiden Mikrofone gleichermaßen beaufschlagenden Geräuschanteile — insbesondere Nebengeräusche — eliminiert, so daß die resultierenden Restsignale ohne weiteres als Sprachbefehle aufgefaßt werden können. Durch den Beginn des nach der Subtraktion verbleibenden Sprachsignals soll z. B. eine Umschaltung eines Freisprechetelefons vom Empfangs- auf den Sendebetrieb nur bei echtem Bedarf erfolgen.

Um den Anteil an Sprachschallbeaufschlagung des Nebengeräuschmikrofons beim Sprechen der Bedienperson möglichst gering zu halten, muß dieses in einiger Entfernung vom Mund der Bedienperson angeordnet sein. Daraus resultiert aber, daß der Nebengeräuschpegel am Nebengeräuschmikrofon nicht unbedingt dem am Sprachmikrofon entspricht, so daß auch ohne Sprachbeaufschlagung des letzteren das Ergebnis der Subtraktion der Signale beider stets auf Sendebereitschaft geschalteter Mikrofone nicht gleich Null sein muß. Daher ist eine Fehlsteuerung auch hier nicht auszuschließen.

Es sind ferner andere Vorrichtungen mit vergleichbarer Funktion bekannt, die je ein einzelnes Sprachmikrofon nur unter bestimmten Bedingungen auf Sendebereitschaft umschalten.

In allgemein bekannter Weise geschieht das, indem der Mikrofonbenutzer von Hand eine Sendetaste betätigt.

Aus der GB 20 03 002 A ist eine Vorrichtung bekannt, die einen Senderschalter bei plötzlich ansteigendem Mikrofon-Eingangsspegel automatisch schließt.

Aus der DE-OS 34 01 883 geht schließlich der Vorschlag hervor, ein Mikrofon mittels eines Annäherungssensors — der die Annäherung des Kopfes der Bedienperson an das Mikrofon zum Sprechen erfaßt — auf Sendebereitschaft umzuschalten.

Bei all diesen Vorrichtungen wird immer lediglich ein (Sprach-)Mikrofon verwendet.

Ferner ist aus der US-PS 33 83 466 ein Verfahren zur Spracherkennung bekannt, bei dem neben Meßfühlern für die Lippenbewegung einer Bedienperson und die Luftströmung zwischen ihren Lippen ein erstes Mikrofon am Kehlkopf und ein zweites in der Nase der Bedienperson angebracht sind. Durch die beiden komplizierten Meßfühler und das Mikrofon in der Nase wird dieses Verfahren für den breiten Einsatz an Arbeitsplätzen oder in Kraftfahrzeugen ungeeignet.

Die Erfindung hat die Aufgabe, das Verfahren nach dem Oberbegriff des Patentanspruchs 1 so zu verbessern, daß eine Deutung von reinen Nebengeräuschen als Sprachbefehlsbeginn sicher ausgeschlossen werden kann, und eine Vorrichtung zur Durchführung des erfindungsgemäßen Verfahrens anzugeben.

Diese Aufgabe wird erfindungsgemäß mit den kennzeichnenden Merkmalen des Verfahrensanspruchs 1 bzw. der Vorrichtung nach Anspruch 2 gelöst.

Die Unteransprüche offenbaren mit ihren kennzeichnenden Merkmalen vorteilhafte Ausbildungen der erfindungsgemäßen Vorrichtung.

Da beim Sprechen erzeugter Körperschall mit äußerst geringem Zeitverzug am Kopf oder Hals, speziell am Kehlkopf der Bedienperson, abgenommen werden kann, eignet sich ein durch ein Körperschallmikrofon erzeugtes Signal sehr gut dazu, den Beginn eines Sprachbefehls zu erfassen. Das eigentliche Sprachmikrofon, welches auch mit Nebengeräuschen beaufschlagt wird, kann durch das Körperschallsignal genau zum Zeitpunkt des Sprachbefehlsbeginns ohne Fehlermöglichkeit elektrisch zur Sprachsignalübertragung eingeschaltet werden. Die Bedienperson braucht dabei keinerlei willkürliche Hand- oder Kopfbewegung zu machen, weil nach einer Ausgestaltung der Vorrichtung beide Mikrofone an geeigneten Geschirren getragen werden können. Sie ist auch nicht an irgendwelchen Bewegungsabläufen gehindert, vor allem dann nicht, wenn ferner noch ein tragbarer Sender vorgesehen ist, der die Sprachbefehle drahtlos zu einem entsprechenden Empfänger überträgt.

Die elektrischen Signale des Körperschallmikrofons sind zwar allein nicht hinreichend für die Spracherkennung, weil sie z. B. keine Formanten und Nasallaute wiedergeben können, aber sie können das übertragene Sprachmuster des Sprachmikrofons weiter verbessern, wenn sie mit dessen elektrischen Signalen in geeigneter Weise verknüpft werden, z. B. durch phasengleiche Addition.

Es ist ferner vorteilhaft, daß das Sprachmikrofon immer wieder automatisch abgeschaltet wird, wenn Sprechpausen entstehen. Die Einschaltdauer kann auch automatisch auf die zulässige Wortempfangsdauer der Spracherkennung beschränkt werden. Im letzteren Fall erfolgt eine Signalisierung an die Bedienperson.

Ein Ausführungsbeispiel für eine erfindungsgemäße Vorrichtung ist in der Zeichnung dargestellt und wird im folgenden näher beschrieben. Es zeigt

Fig. 1 eine schematische Darstellung der Vorrichtung und

Fig. 2 die Verschaltung der Mikrofone und der Triggerschaltung in einem Blockschaltbild.

An einem Halsgeschirr 1 ist ein an den Hals einer nicht dargestellten Bedienperson anzulegendes erstes Mikrofon 2 — als Doppel-Kehlkopfmikrofon ausgeführt — befestigt. An einem Kopfgeschirr 3, das zu einem Kopfhörer 4 gehört, ist ein auf den Mund der Bedienperson auszurichtendes zweites Mikrofon 5 befestigt. Vom ersten Mikrofon 2 führt eine Signalleitung 6, in die ein manuell zu betätigender Unterbrecherschalter 6.1 eingeschleift ist, zum Kopfgeschirr 3 und wird von dort aus gemeinsam mit einer Signalleitung des zweiten Mikrofons 5 in einem flexiblen Kabel 7 über eine mehrpolige Steckverbindung 8 in ein erstes Gehäuse 9.1 eines Gehäuses 9 eingeführt.

In dem ersten Gehäuse 9.1 ist eine — hier nur schematisch mit einem Transistorsymbol dargestellte — Triggerschaltung 10 angeordnet, welcher die elektrischen Signale der beiden Mikrofone 2 und 5 zugeführt werden. Schließlich ist noch ein Signalgeber 10.3 angeordnet, der von einem in der Triggerschaltung 10 vorgesehenen Zeitglied aktiviert werden kann und der der Bedienperson ein optisches oder akustisches Signal bei Überschreitung der maximalen Wortempfangsdauer einer Spracherkennungsschaltung 14 gibt.

In einem zweiten Gehäuse 9.2 des Gehäuses 9 sind ein Sender 11 und ein elektrischer Energiespeicher 12 angeordnet. Der Sender 11 korrespondiert mit einem Empfänger 13 der Spracherkennungsschaltung 14. An dem zweiten Gehäuse 9.2 kann vorteilhaft ein Trage-

griff oder -gurt 15 angebracht sein.

In dem stark vereinfachenden Blockschaltbild der Fig. 2 sind die den bereits erwähnten Schaltelementen der Vorrichtung entsprechenden Schaltsymbole mit gleichen Bezugsziffern versehen.

Die Triggerschaltung 10 besteht im wesentlichen aus einem Schwellwertschalter 10.1 mit z. B. an einem Potentiometer einstellbarem Referenzwert und aus einem von dem Schwellwertschalter 10.1 steuerbaren Schließerschalter 10.2, der hier in den Signalweg des zweiten Mikrofons 5 eingeschleift ist. Der Schließerschalter 10.2 enthält vorteilhaft ein Zeitglied T , welches nach Ablauf der maximalen Wortempfangsdauer der Spracherkennungsschaltung 14 den Schalter 10.2 öffnet und den Signalgeber 10.3 zu dem erwähnten Zweck aktiviert.

Der Sender 11 kann nach beliebigen Prinzipien der drahtlosen Signalübertragung arbeiten. Als Beispiel ist eine Infrarotdiode eingezeichnet.

Ferner sind im ersten Gehäuseteil 9.1 noch ein Verstärker 16 für die Signale des ersten Mikrofons 2 und ein Verstärker 17 für die Signale des zweiten Mikrofons 5 angeordnet.

Bei Beaufschlagung des ersten Mikrofons 2 mit Körperschall wird nach dem Ansprechen des Schwellwertschalters 10.1 der Schalter 10.2 geschlossen bzw. durchlässig für die verstärkten und gefilterten Signale des zweiten Mikrofons 5.

Sollte die Bedienperson sich räuspern oder husten müssen, kann sie mit dem Unterbrecherschalter ("Räuspertaste") 6.1 eine unerwünschte Signalübertragung verhindern.

Schließlich kann ein akustisches Signal, das die Überschreitung der maximalen Wortempfangsdauer markiert, der Bedienperson vorteilhaft auch über den Kopfhörer 4 statt über den direkt abstrahlenden Signalgeber 10.3 zugeführt werden.

Hierzu 2 Blatt Zeichnungen

40

45

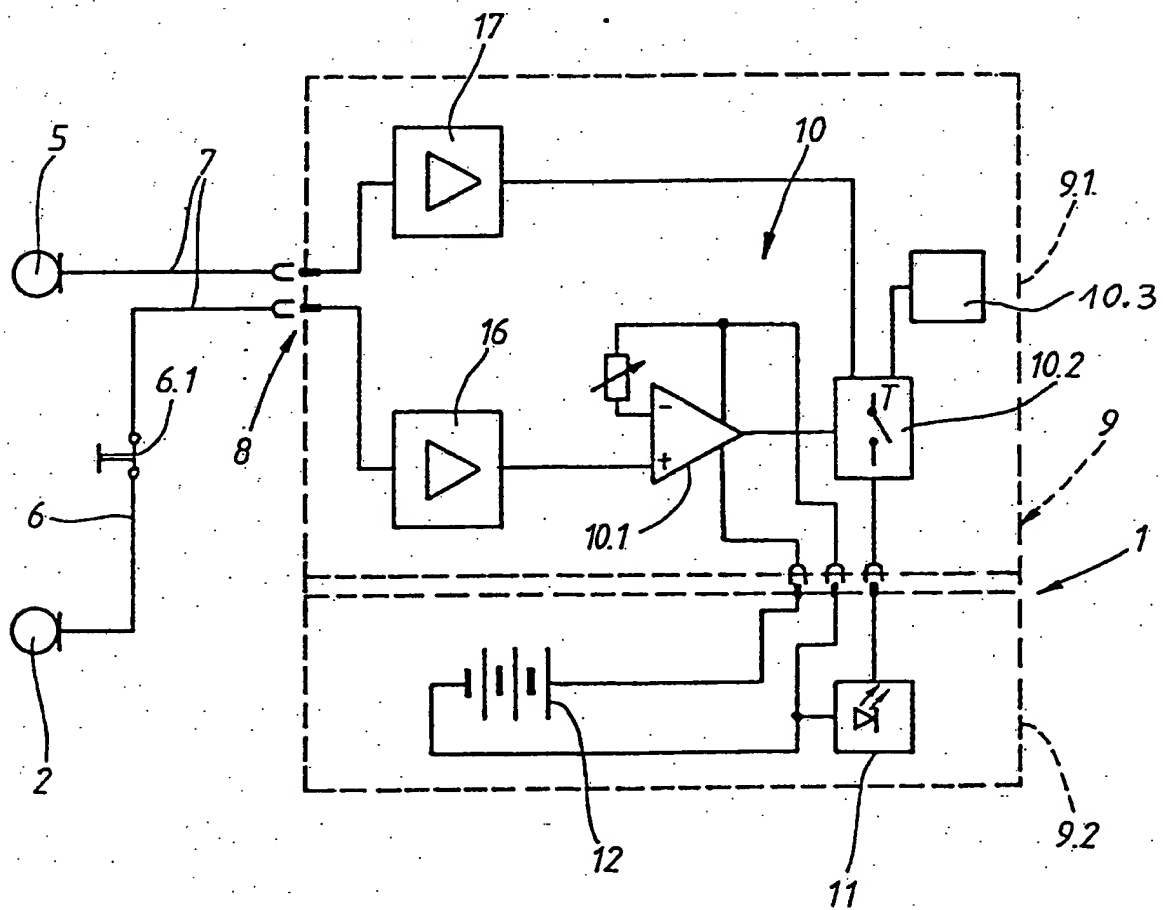
50

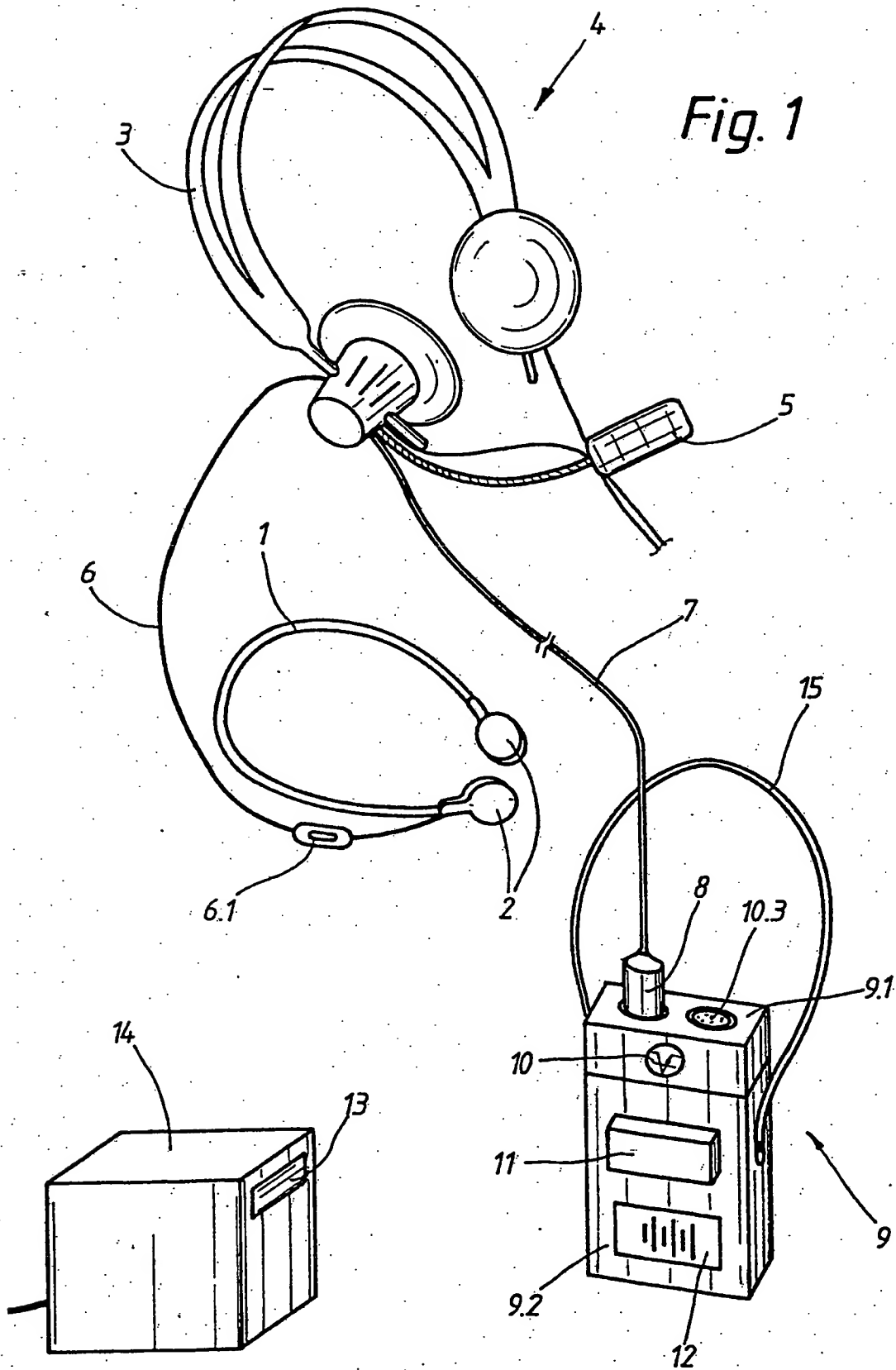
55

60

65

Fig. 2





Increasing reliability of speech controls - triggering mouth microphone from signal originating from microphone responding to body sound of operator

Patent Number : DE3742929

International patents classification : G10L-003/00 G10L-007/08 H03G-003/20 H04B-003/00 H04B-015/02 H04Q-009/00

• Abstract :

DE3742929 C A first microphone (2) is fed with body sound forming speech by an operator. A second microphone (5) is spatially separated from the mouth of the operator so as to receive the speech sounds and also ambient noise. An electric circuit creates a set relation between the signals from the two microphones and feeds an output to a speech recognition circuit. It comprises a trigger circuit (10) making a switch relation between the two microphones so as feed a signal from at least the second microphone to the speech recognition circuit only when simultaneous signals from the first microphone appear at the trigger.

The first microphone may be a larynx-microphone attached to the throat of the operator. Both microphones may be suspended from harnesses.

USE - Ensures recognition of beginning of spoken command in speech recognition circuit. (2/2)

US4901354 A The reliability of voice control systems is improved by reliably detecting the beginning of voice commands in a voice recognition circuit. Sound generated during speaking is applied to a first microphone with little ambient noise present and a second microphone some distance away. Misinterpretation of a pure ambient noise as voice command is prevented by signals generated in both microphones being electrically subtracted. Contact measurable sound generated during the speaking of an operating person is applied to the first microphone only direct contact. The second microphone is directed towards the mouth of the operating person. As a result, sound generated during speaking is also applied. This microphone is electrically switched to the voice recognition circuit only when the first microphone generates signals from the contact measurable sound.

USE/ADVANTAGE - At work station or in motor vehicle. Better rejection of ambient noise allowing speech control of systems where operator is manually restricted.

(6pp)

• Publication data :

Patent Family : DE3742929 C 19880929 DW1988-39 6p * AP:
1987DE-3742929 19871218
GB2211336 A 19890628 DW1989-26 AP: 1988GB-0029362
19881216
FR2625020 A 19890625 DW1989-32
JP01194797 A 19890804 DW1989-37 AP: 1988JP-0304328
19881202
US4901354 A 19900213 DW1990-13 6p AP: 1988US-0279814
19881205
Priority N° : 1987DE-3742929 19871218
Covered countries : 5
Publications count : 5
Additional words : WORK STATION MOTOR VEHICLE

• Patentee & Inventor(s) :

Patent assignee : (DAIM) DAIMLER-BENZ AG
Inventor(s) : GOLLMAR K; JUENEMANN G

• Accession codes :

Accession N° : 1988-271995 [39]
Sec. Acc. n° non-CPI : N1988-206594

• Derwent codes :

Manual code : EPI: W04-V01 X22-L
Derwent Classes : P86 W04 X22

• Update codes :

Basic update code :1988-39
Equiv. update code :1989-26; 1989-32;
1989-37; 1990-13

THIS PAGE BLANK (USPTO)